Quantifying Human Behaviour in a Retail Environment

Amr Dahawi

A thesis submitted in partial fulfilment of the requirements for the Degree of Master of Computer Science and Software Engineering in the University of Canterbury

Supervised by Richard Green

2016

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

ABSTRACT

Robustly quantifying human behaviour in a retail environment raises research challenges around accurately and reliably recognising motion, age, gender, repeat customers and product acquisition in such unconstrained conditions.

The motivation for this research is that computer vision can be used in the retail refrigeration industry to provide the shop/product owners with information about their clients, products, sales, stock levels and can also help with understanding the customers' needs and psychology.

This proposed method improves the accuracy of traditional face detection and recognition using depth information, in uncontrolled lighting environments and where the orientation of faces are not only front facing.

Further proposed algorithms are tested on product recognition from a retail refrigeration unit in a retail setting. These proposed methods adapt Hue manifold, Haar cascade classifiers, SIFT and Local Binary Patterns Histograms.

The face detection results of 96% recall and 100% precision together with face recognition results of 85% recall and 97%, indicates that the proposed method may be useful for improving face recognition in variable lighting environments where people do not stop moving and are not always facing the camera.

The product recognition results shows a 58% precision in a controlled lighting environment comparing to 38% precision in an uncontrolled lighting environment.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Richard Green, Associate Professor, Department of Computer Science and Software Engineering, University of Canterbury, for his advice and guidance during this research. I have learned a lot during this research and benefited from his knowledge and expertise in Computer Vision.

Finally, I would like to thank my wife and parents for their help and support, and would have not been able to achieve this without them.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

The proposed face and object recognition methods are two different proposed approaches that we will be combining together in this research. The motivation for this research is that computer vision can be used in the retail refrigeration industry to provide the shop/product owners with information about their clients, products, sales, stock levels and can also help with understanding the customers' needs and psychology. But robustly quantifying human behaviour in such retail environments raises research challenges around accurately and reliably recognising motion, age, gender, repeat customers and product acquisition in such unconstrained conditions.

Face recognition is used in many applications and as a result can replace many of the expensive manual authentication methods used traditionally, such as security and surveillance systems.

In security Systems it has the potential to replace traditional verification methods like access cards, passwords and pin numbers. Face recognition is more commonly found in places with high security requirements like airports and other boarders [1, 2].

Because CCTVs are now so commonly used in shops, airports and streets, face recognition can be used in real time to look for and recognize wanted/known persons of interest [3].

This research investigates how the tracking of a person using depth information can improve the process of face detection/location and recognition.

There are two main factors that affect the appearance of faces, which can be categorized [4] as:

A. Intrinsic category

This refers to the physical features of the face, where there are two factors [5] here:

a. Intrapersonal

This refers to different facail features of the same person which occurs over the years due to age, facial hair or even the use of glasses or lenses.

b. Interpersonal

This refers to differences in facial features due to colour, ethnicity or gender.

B. Extrinsic category

a. This can also be any external factors such as lighting and face orientation (frontal, side)

Figure 1 Face Recognition

Robust accurate object recognition is one the most challenging fields in computer vision and pattern analysis research.

This challenge is emphasised by the contrasting fact that a human can detect and recognize objects such as cars and animals easily and with high accuracy.

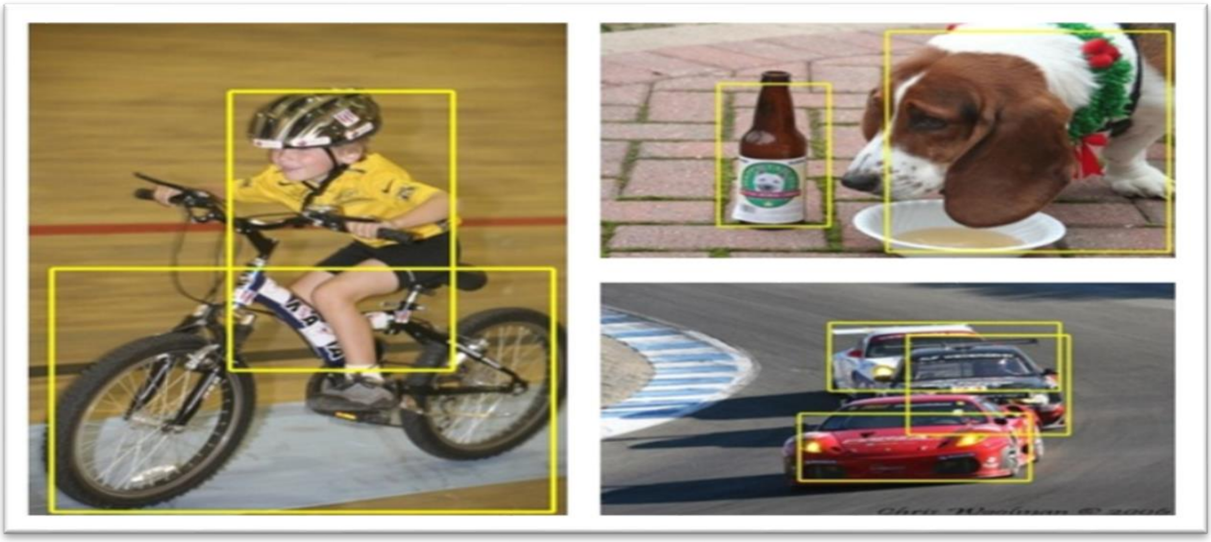


Figure 2 Object detection

1.1 The Problem and Solution

The principle challenge for both face and object recognition is achieving sufficient accuracy. Unfortunately many prior research approaches achieve high levels of accuracy as a result of running the experiments in constrained controlled environments.

Our goal in this research is to achieve high accuracy results in an uncontrolled environment.

A camera will be installed in a retail environment, there will be no control over face orientation, and there will also be no control of how close or far the faces are. Varying illumination levels is also another factor, including images being captured in a natural lighting (sun light) in the morning and indoor lighting at night.

Two types of cameras will be installed in a refrigerator, a Kinect 2 camera will be installed on the top of the fridge and small web cameras will be installed in the door panel.

1.2 Objectives

The goal is to improve the human and object identification and overcome some of the limitations in prior research.

The following are the project objectives and methodologies:

1. Tracking of Multiple Customers' locations

A 3D tracking technique based on depth information will be used to track customers who move in front of the cabinets/fridges in retail environment. The proposed approach is to use depth information to calculate the distance between the customers and the fridges.

4 pre-defined range categories will be used:

long-range, mid-range, close and action.

Converting the depth information to real word coordinates will Identify the closest range a customer gets to (in meters) before moving away from the fridge and also the pattern and the directions of customers motions in front of the camera.

The Tracking algorithm is designed to track up to 6 people at any time, this number is based on the area that can be covered by the camera and also taking in account Personal space.

2. Face (Customer) Recognition

Local Binary Patterns Histograms Method, skin colour detection and depth information will be used to detect and recognise faces which will then be used to differentiate between new and existing enrolments. The reason for combining these three methods is to overcome the limitation of only using the Local Binary Patterns Histograms method which can be significantly degraded by lighting conditions, camera position and angle.

3. Object Reconstruction and Recognition

Proposed object reconstruction and recognition algorithms will be used to reconstruct objects from images to identify products. The proposed algorithm will use images from web cameras on every shelf to capture images of the shelves.

The proposed algorithm will be tested by capturing images of products with different angles and in different lighting conditions, product location will then be determined and features will be extracted.

Figure 3 Object recognition

4. Planogram Compliance

Detecting objects using appearances, structural and shape features will be used to identify products in the images combined with colour histogram technique and feature extraction for product recognition.

Each Fridge has a planogram that illustrates what the layout should be, and the goal of this research objective is to find out if the physical layout of the products complies with the provided planogram.

This will be achieved by storing the planogram information in a database/file with all the information of the layout (ie. Shelf 1, row 1, Coke) and then recognizing the actual layout of the product and comparing it to the planogram information.

1.3 Thesis Guide

Chapters two and three will describe the different algorithm and techniques available for face, object detection and recognition and look at the advantage and limitations of these algorithms and how they were used in prior research.

Chapters four and five will discuss the proposed methods of this research and their implementations. This will include information about the hardware used in this research.

Chapter Six, will discuss the experiment carried out to test the proposed methods and provide experiment results and analysis. Finally chapter Seven, the conclusion, will provide a summary of the research and future work.

2. Face Recognition

Face recognition consists of two main components, face detection and face recognition. Once a face is detected then it can be recognized (if already known to the system or as a new face) and tracked from one frame to the next.

2.1 Face Detection

Face Detection is the process of locating a face in an image or a scene, the goal is to locate the face only without including anything from the background. The approaches for face detection has been classified into the four main categorise [6, 7], knowledge based, feature invariant based, appearance based and template matching methods as discussed in the following subsections.

2.1.1 Knowledge based method

Knowledge based method is premised on our knowledge of facial features (i.e. two eyes, a nose and mouth in a known configuration). The approach works by setting rules which are used to locate face features based on their distance and intensity. This approach was one of the early algorithms used to only detect frontal faces but it is not efficient and is not rotational invariant. Also it is not easy to set a rule that can be sufficiently generic to give good results and yet only work on specific images. However, a combination of specific and generalized rules can be used in one system, as was demonstrated in [8, 9].

2.1.2 Feature invariant based method

This approach is based on finding invariant features in human faces, where such features can be the eyes, nose, and mouth. These features have often been detected using line or edge detection methods and then used to build a statistical model which is used to find or verify if an area of the image is a face or not. One disadvantage of this method is it can be highly affected by lighting, impacting feature location of the face (as shadows can create fake lines or edges). There have been many different approaches implemented by authors [10] which are based on canny detector, and in [11] where a combination of methods were used including band pass filtering, morphological operations, histograms and classifiers.

Some of other approaches that were used are based on texture like in [12]. Detection based on skin colour [13, 14, 15] characteristics is another approach, although detection using skin colour [16, 17] can cause false detections if the background contains objects with similar colour as skin, such as wood panelling.

Techniques based on eye blinking as implemented by the authors in [18] can also produce good results but some of the limitation of this approach is when involuntary blinking can raise false detection or when it does not occur enough to be detected.

2.1.3 Appearance based method

The appearance based method is based on creating templates that are learnt from the images, this process uses statistical and machine learning algorithms.

Some of these statistical analyses includes Principal Component Analysis (PCA) which was demonstrated in [19]. PCA is used to decompose the vector space into two mutually exclusive and complementary subspaces: the principal subspace (or feature space) and its orthogonal complement.

The subspace represents a set of face patterns. The principal components preserve the major linear correlations in the data and discard the minor ones. Local features of the face are then learnt using multivariate Gaussian and a mixture of Gaussians. Other analysis tools are Linear Discriminant Analysis (LDA) and Factor analysis (FA), which were used in [20].

Another appearance based method is based on Neural Networks and advanced usage of this method was introduced in [21].

Neural Network is another method used for face detection, as face detection can be considered as a two class problem (face and non-face).

Other statistical methods used in face detection include support vector machines [22] and Naive Bayes Classifiers [23].

2.1.4 Template Matching method

Template matching method is based on creating a predefined template for the face, which would hold (usually) the frontal face pattern (or side face profile). Once a template is available it can be used to match against an input image based on the correlation between pattern in the input image and the template.

One of the most popular approaches in Face detection is based on using Haar feature-based cascade classifiers [24, 25]. This approach is based on the following four key concepts:

1. Haar Features

Consist of a rectangular feature consisting of light and dark regions. Using this template instead of individual pixels, also results in speed improvements.

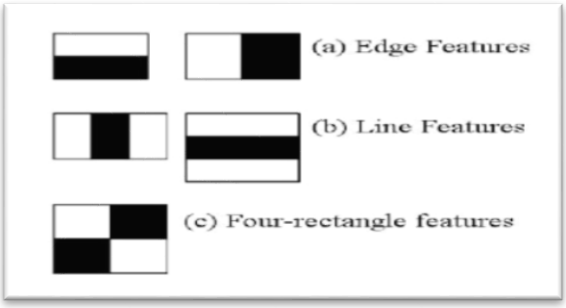


Figure 4 Examples of Haar features

2. Integral Image

Is an intermediate representation for the image at a point, the value of this integral image is the sum of all pixels above and to the left of this point.

$$I(x,y) = \sum_{\substack{x' \le x \\ y' \le y}} i(x',y')$$

3. AdaBoost [25] machine-learning Method

Is used to select the features and to train the classifiers, it also improves the performance of the algorithms as it uses cascaded classifiers.

4. Cascaded Classifier

A combination of weak classifiers combined to form a stronger classifier.

Viola-jones algorithm [23] works by applying filters consisting of adaboost classifiers at different sub-areas of the image (integral Image). When a filter passes an area, the next filter in the chain is applied until the end of the chain and then the area is classified as a 'FACE'. If a filter fails at any point it classifies the area as 'Not Face' and moves on to another area. Filters are ordered by adaboost depending on the filter's importance weight.

Figure 5 Face detection (positive and negative results)

2.2 Face Recognition

After identifying and locating a face, Face Recognition is the second main component, where a face is compared to a database of images to find out if this person/image has already been enrolled.

There are two main categories for face recognition, one is based on holistic methods and the other is features-based.

2.2.1 Holistic based method

Linear Discriminant Analysis (LDA) is a classification method originally developed in 1936 by R.A. Fisher. It is a simple and robust algorithm that produces accurate models. LDA is based on searching for a linear combination of variables (predicators) that best separates two classes. It attempts to express one dependent variable as a linear combination of other features of measurements. LDA [28] is an enhancement to PCA as it minimizes the scatter between images of the same class and maximizes the scatter between different class images.

There are two measures for all the samples of all classes:

Within-class scatter matrix

$$S_w = \sum_{j=1}^{c} \sum_{i=1}^{N_j} (\mathbf{x}_i^j - \mu_j)(\mathbf{x}_i^j - \mu_j)^T,$$

Where x^j_i is the ith sample of class j, U_j is the mean of class j, c is the number of classes, and N_j the number of samples in class j;

And, between-class scatter matrix

$$S_b = \sum_{j=1}^{c} (\mu_j - \mu)(\mu_j - \mu)^T,$$

where U is the mean of all classes.

Two- Dimensional Principal Component analysis (2DPCA), 2DPCA [29] technique is developed for image representation and for image feature extraction.

It is based on 2D matrices, in contrast to the image covariance matrix which is constructed directly by using the original image matrix, the size of the image covariance matrix using 2DPCA is much smaller.

2DPCA have lots of advantages like it is easier to evaluate the covariance matrix accurately and it also takes less time to determine the corresponding Eigenvector [30].

Eiginfaces method [26] is considered to follow the holistic approach, EigenFaces (as shown in figure 6) are extracted from original image by a mathematical tool called Principal Component Analysis (PCA) [27]. PCA produces eginvectors of the face images which hold specific characteristic features. Original images can be reconstructed by combining these eigenfaces.

Face recognition works by computing the Euclidean distance between the input image and the training image that is most similar to it.

Euclidean distance Equation for point1 (P1) and point 2 (P2) in 2D:

d12=sqrt(dx2+dy2)

where dx = x2-x1, and dy = y2-y1

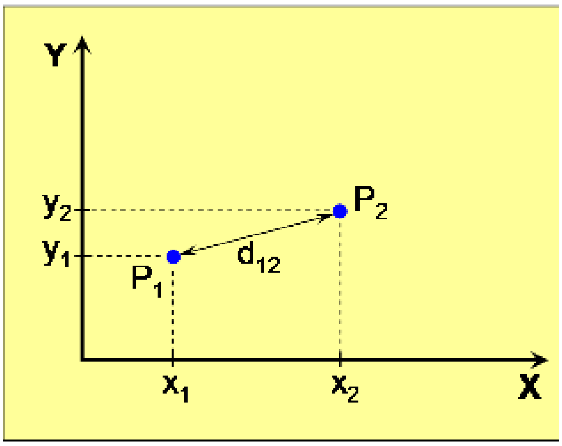


Figure 6 Euclidean distance in 2D

Limitations of Eigenface is that it can be highly affected by lighting conditions, distance of the camera from the face, angle of the camera - but these can be minimized by constraining the face capture environment such as ensuring that the camera is at the same level as the face which is always in a frontal orientation.

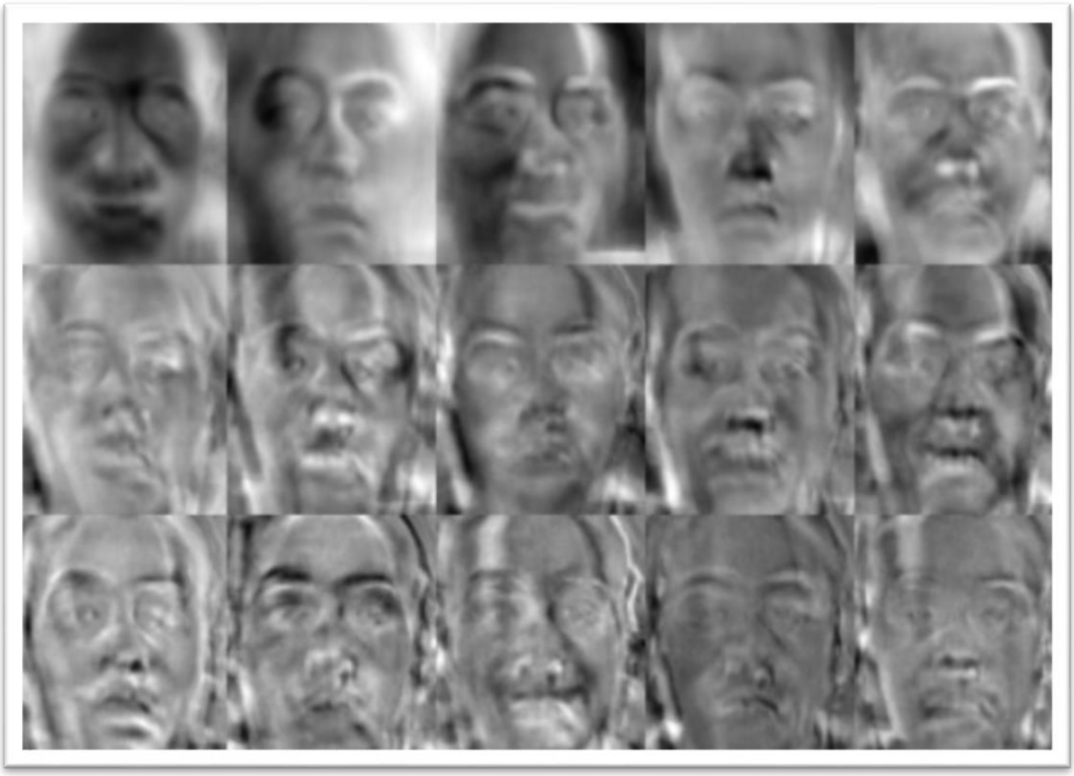


Figure 7 EigenFaces

2.2.2 Feature based method

Feature-based algorithms use local features such as the eyes, nose, and mouth to supply their locations and local statistics into a structural classifier. One of the advantages of using this is to avoid the high-dimensionality of the data as it only describes local regions of the images.

Gabor Wavelets, was proposed by Hungarian-born engineer Dennis Gabor in 1946. It is manly used now for feature extraction and face recognition [31]. It is a filtering device that transforms facial images into wavelets. One of its advantages is that it is unaffected by variations of light, facial expressions or poses. Gabor wavelets are represented by 2-D plane waves in the spatial domain. Wavelets can be located somewhere between the space and the frequency domain.

Discrete Cosine Transform, DCT [32] is an approach to compress images, it works by removing all information that is not used. It is similar to the discrete Fourier transform, where it transforms a signal or image from the spatial domain to the frequency domain.

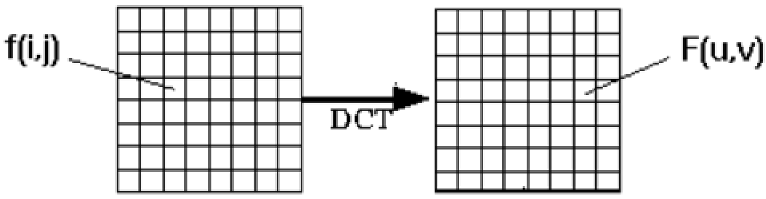


Figure 8 DCT

The General Equation for 2D DCT is:

$$F(u,v) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \Lambda(i).\Lambda(j).cos\left[\frac{\pi.u}{2.N}(2i+1)\right] cos\left[\frac{\pi.v}{2.M}(2j+1)\right].f(i,j)$$

Local Binary Patterns [33, 34, 35], LBP considers the surrounding points of a central point and tests these points to find out if these points are greater or less that the central point.

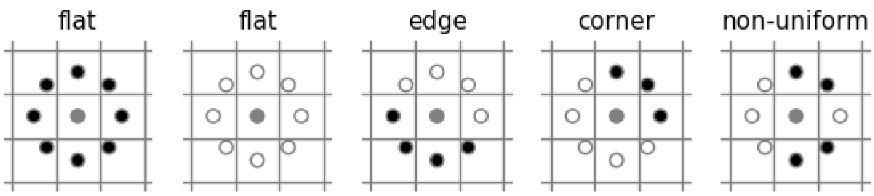


Figure 9 Surrounding points

As shown in figure 9, the black points are less intense that the central pixel while the white ones are more intense.

The LBP operator can vary from a 3x3 neighbourhood to any size.

When all the surrounding pixels are black or white, then this would be considered a flat region of the image (featureless). A group of successive black or white pixels are considered “uniform” patterns (ie. Corners or edges). When the pixels keep alternating between black and white, this is considered a “non-uniform” pattern.

While computing the LBP histogram, all the non-uniform patterns are assigned a single bin, while every uniform pattern is assigned a different bin

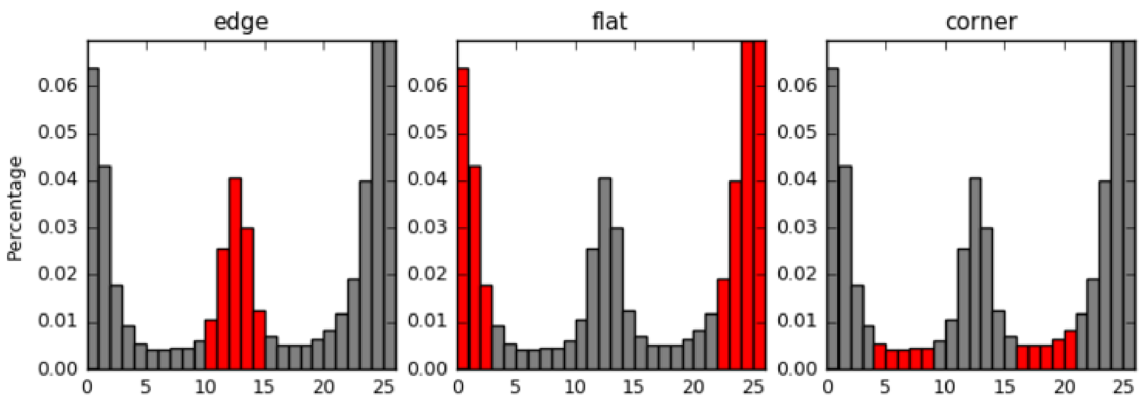


Figure 10 LBP Histogram

3. Object Recognition

Object recognition is similar to face recognition in that it consists of two components; one is object segmentation and the other is object recognition.

Figure 11 Object detection and recognition

3.1 Object Detection

The purpose of object detection is to locate/identify one or more objects in a scene/image.

There has been extensive prior research done on this area as it is used in many domains, including techniques such as:

Statistical template matching with sliding window [36], is one of the non-recursive techniques that uses a sliding window approach.

This approach applies a classifier function to all the sub images/windows. Once the function is applied, the classification with the maximum score would then indicate the presence of an object in the region tested.

A low resolution image (320 x 240) contains more than 1 billion sub images/window, which mean that this approach can be computationally too expensive. This has been addressed by some other methods as in [37].

Another approach is based on detection by parts using appearances, structural and shape features [38]. Structural constraints are represented with a constellation model [39], where all object parts are fully connected (as graphs), and all parts are dependent on each other for the same object.

Appearance based detection using HOG [40] is a region descriptor which works by calculating the spatial distribution of gradient orientations and is partially invariant to illumination and rotations.

Some other recursive methods [41] such as Frame Differencing, optical flow and background subtraction can be used (see Table 1)

Frame differencing, is based on calculation the difference of two consecutive images. The result image will indicate the presence of any moving objects. This approach has an issue which is referred to as the ghost effect that is caused by moving objects. This has been researched and improved in [42].

Optical Flow [43], is the distribution of apparent velocities of movement of features in an image. Optical flow arises from camera/object movements. It can give important information about the spatial arrangement of the objects viewed and complete movement information. Discontinuities in the optical flow can help in segmenting images into regions that correspond to different objects.

Background subtraction, is based on subtracting each frame from a background reference image, which is created via background modelling.

Background subtraction will highlight moving objects in the current frame. This basic algorithm has some disadvantages as it is not robust to dynamic backgrounds. There are two main approaches in Background subtractions:

- Recursive algorithm:
    - With this technique there is no need to maintain a buffer for background, the background model is recursively updated after each frame. This adaptive background technique includes various methods such as approximate median, adaptive background, Gaussian of mixture.

- Non-Recursive algorithm:
    - This technique unlike recursive algorithm needs to store the buffer n frames and uses a sliding window approach for background estimation.

      Non-recursive algorithms are highly adaptive but in the other hand might require a significant storage for the buffer.

Methods		Accuracy	Computational Time	Comments
Background Subtraction	Gaussian Of Mixture	Moderate	Moderate	+ Low memory requirement - It does not cope with multimodal background
	Approximate Median	Low to Moderate	Moderate	+ It does not require sub sampling of frames for creating an adequate background model. - It computation requires a buffer with the recent pixel values
Optical Flow		Moderate	High	+ It can produce the complete movement information

			- Require Large amount of calculation
Frame Differencing	High	Low to Moderate	+ Easiest Method. Perform well for static background. - It requires a background without moving objects

Table 1 Recursive Methods Comparison

3.2 Object Recognition

Object recognition is the second phase of object identification and the purpose of this is to try and match/categorize the objects detected to a known object/category (ie, car, plane, tree, etc…).

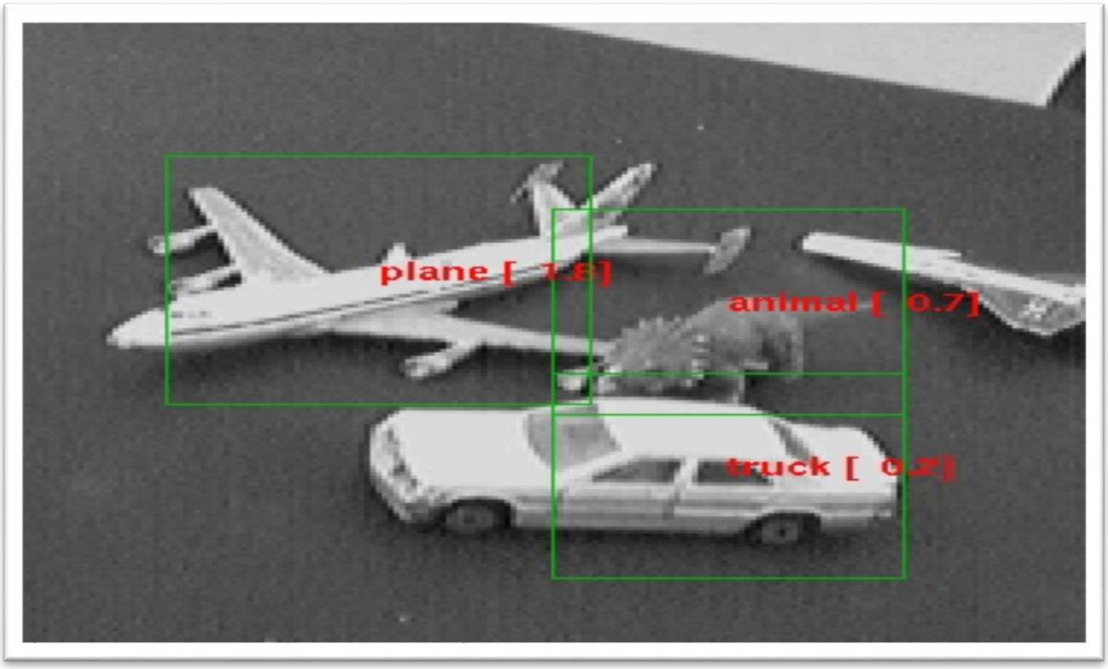


Figure 12 Objects Recognition

Some simple methods can be used, such as colour histograms [44] which compute the colour histogram of different images and compare them. Colour co-occurrence histogram is another approach which is very similar to colour histogram but it also includes the distance between two pixels which gives the histogram additional geometric information about the object. Those two methods have the limitation of being affected significantly by illumination variations.

Object recognition can also be achieved using feature extraction methods such as SIFT [45, 46], SURF [47] and then features can be matched using algorithms like Brute-Force matcher [48] or FLANN matcher [49].

One of the recent approaches that has been researched and implement is Discriminatively Trained Part-Based Models (DPM) [52]. This approach is based on three characteristics:

1. Strong low-level features based on histogram of oriented gradients (HOG)
2. Efficient matching algorithms for deformable part-based model (pictorial structures)
3. Discriminative learning with latent variables (latent SVM)

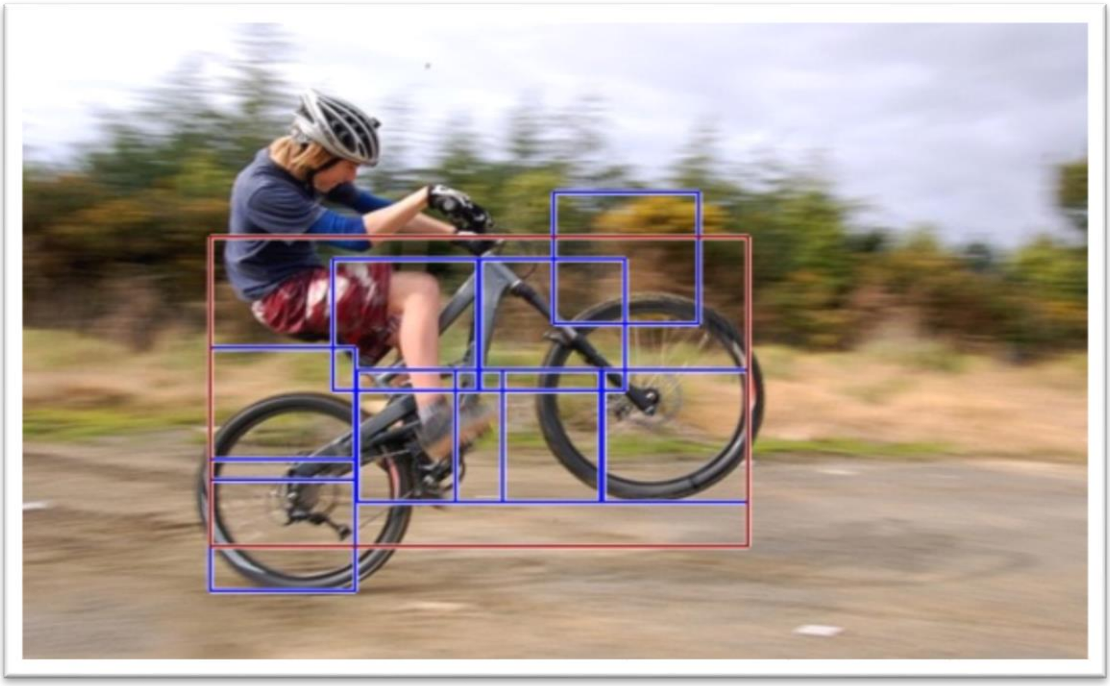


Figure 13 Object recognition using DPM

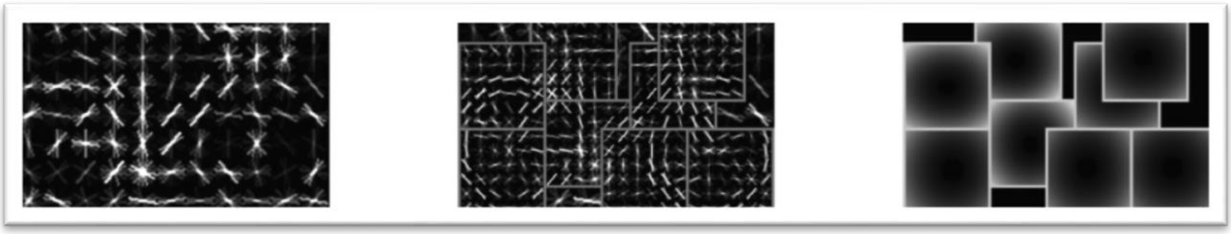


Figure 14 DPM - part-based learning

3.3 Summary

There has been a large body of research into object recognition. But in our research, the challenge is to achieve accurate recognition without having to train a model using thousands and thousands of images as is necessary to take advantage of recent research using Convolutional Neural Networks (CNN or deep learning).

4. Proposed Methods for Face Recognition

The Proposed method was implemented on a Windows 8 64bit laptop.

Every captured image consisted of a pair of frames, one is the colour frame and the other is the depth frame.

Every time an image (pair of frames) is captured, it gets pre-processed and saved in memory. The reason for saving these frames in memory is to be able to run in a multi-threaded mode, where the main thread is only capturing and saving the frames.

The secondary thread retrieves the captured frames and applies the proposed method. Using Multi-threading improves the overall performance of the proposed algorithm which leads to better results.

4.1 Colour and Depth frames Pre-Processing

Colour frame has all the required information such as frame size and raw image content data. The frame content is converted an n-dimensional array to represent the image.

The image is then processed further using tracking, where if a human is detected then additional information is saved with the frame, including head position and tracking ids of the human bodies.

Other information is processed from the depth frame to determine the location of the user.

All this logged information can be processed later on to understand behaviour and patterns of customers in regards to conversion rates.

Depth information is captured up to only ~3 meters form the camera.

Figure 15 Tracked Person

Figure 16 Depth Frame

4.2 Processing Colour frames for Face detection

While the frames are being captured, another thread is running in parallel to process the captured frames. This thread will process a frame (pair) at a time.

4.2.1 Skin detection

Is applied to the colour frame, where the traditional skin detection approach is to find all pixels in the frame that fall within a range of colour (skin hue manifold) using HSV space. The improvement introduced in this proposed method is to use a combination of the skin detection and the head position to eliminate any false positive detected skin area.

Below is the proposed algorithm for the improved/combined skin detection and head position:

1. Copy the colour image and convert it to HSV
2. Smooth image using Gaussian filter
3. Loop through each pixel in the frame staring from (0,0) position
4. Check if the current pixel falls in the range of skin colour (HSV space)
5. Check if the pixel falls in an area within a certain (configurable) distance from any of the provided head positions
6. If conditions 4 and 5 are met, then keep the colour of the pixel as is, otherwise change it to a black pixel
7. Change the image back to BGR colour space
8. Change the image to black and white

9. Apply a segmentation operation using binary thresholding

10. Apply 3 morphological operations to the image, Dilate, Close, Dilate.

11. Smooth the image using a median filter.

12. Mask the copied image onto the original frame.

The final (copied) image will consist of white and black pixels, where the white pixels represent the Face area in the image.

This will change the original image to only have visible face.

Figure 17 Original Image

Figure 18 HSV Image

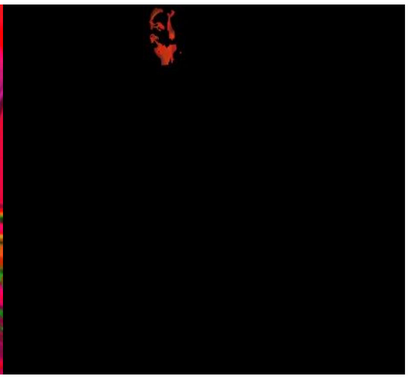


Figure 19 Detected Skin near Head position

Figure 20 BGR Conversion

Figure 21 Grey conversion and Binary Threshold

Figure 22 morphology operations

Figure 23 median filter

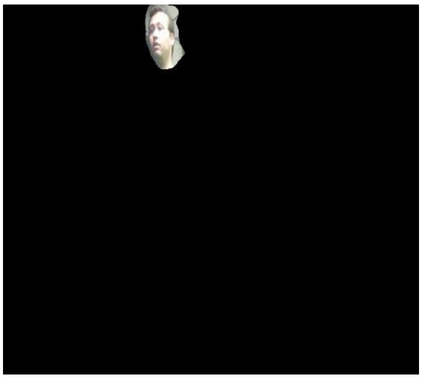


Figure 24 Original image masking

4.2.2 Face detection using Haar Cascade Classifiers

For improved accuracy, another method of face detection will be applied using the image resulted from the skin detection. This it to make sure that only the face is being extracted and any other parts of the background is being ignored/discarded.

Below is the complete process of the Haar cascade classifiers using the resulted image from the skin detection:

1. Convert image to grayscale
2. Apply Histogram equalization method to standardize the brightness and contrast.
3. Apply classifiers to detect face, this process work by:
    a. Create integral images
    b. Applying filters consisting of adaboost classifiers at different sub-areas of the image (integral Image), When a filter passes an area, the next filter in the chain is applied until the end of the chain and the area is marker as a 'FACE'.
    c. If a filter fails at any point it marks the area as 'Not Face' and moves on to another area.
    d. Filters are ordered by adaboost depending on the filters importance weight.
4. List of faces are returned, with location information of each face
5. For each face, match the face to a tracking id.
    a. This is done by finding the centre point of the detected face and for each head position of each tracked person calculate the distance between these 2 points. If the distance falls within a (configurable) range then the face gets assigned a tracking number.

b. Equation for calculating the distance between the 2 centre points is:

   $X = (face.x - head\ position.x)^2$

   $Y = (face.y - head\ position.y)^2$

   $Distance = \sqrt{(X - Y)^2}$

6. Add the face image and its tracking number to a vector to be processed later using face recognition

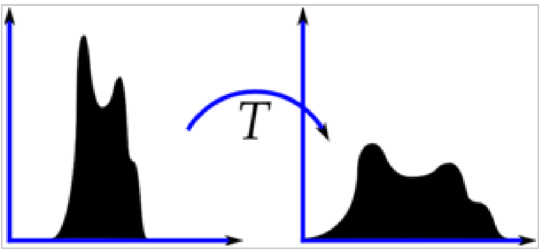


Figure 25 Histogram Equalization

Figure 26 Face Detected using Classifiers

4.3 Face Recognition

The used method for face recognition is based on Local Binary Pattern Histogram [50, 51] method combine with the tracking information provided.

The 2 main reasons for choosing LBP method is that it does not require to look at the whole image as a high-dimensional vector, instead it only needs to describe the local features of the image/face.

The second reason is that a LBP model can be updated without having to retrain the model with all the images.

The LBP works is by dividing the image into n local regions and extract histogram from each image and then concatenate the local histograms which will give us the Local Binary Patterns Histograms.

LBP operator can be represented as:

$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c)$$

With (x_c , y_c) is the central pixel with intensity i_c , and i_n which is the intensity of the neighbour pixel. *s* is the sign function,

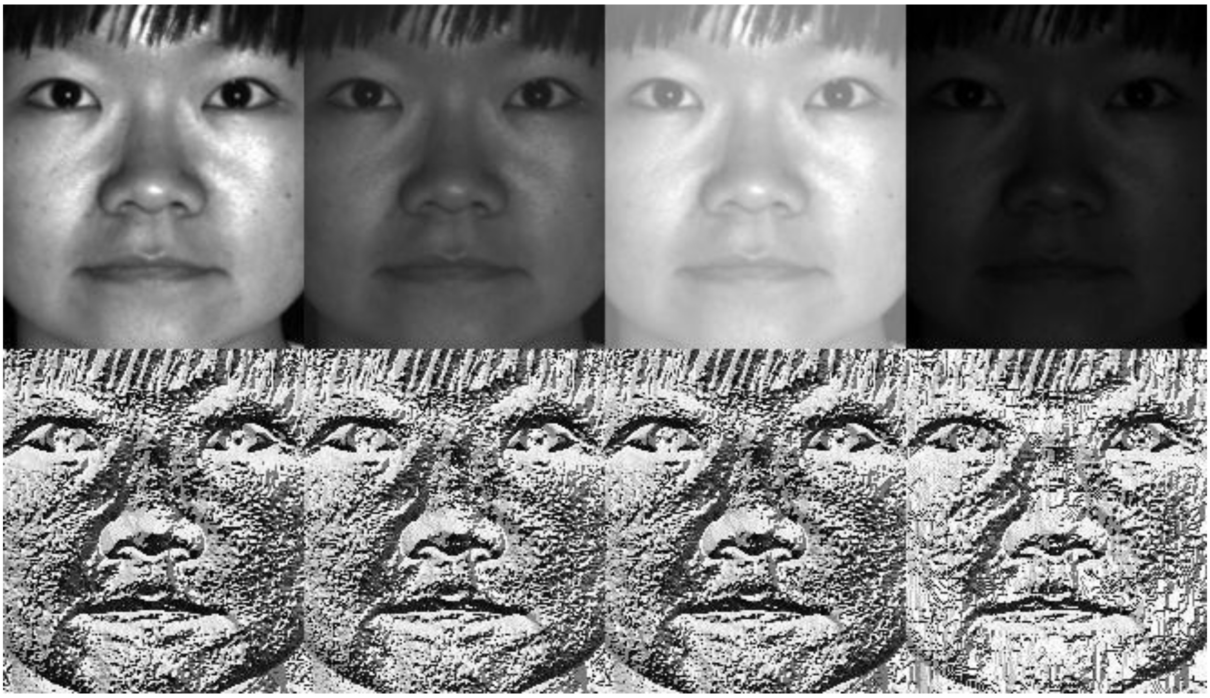
$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases}$$

The position of the neighbour (x_p , y_p), p ε P is:

$$\begin{aligned} x_p = & \quad x_c + R\cos(\frac{2\pi p}{P}) \\ y_p = & \quad y_c - R\sin(\frac{2\pi p}{P}) \end{aligned}$$

Where R is the radius of the circle (number of the neighbours) and p is the number of sample points.

LBP is robust against grey scale transformations.



Below is the complete process of face recognition using LBPH combined with the provided tracking information:

1. Load LBPH model if exists
2. For each new face, check if this has been recognized by the model.
3. If recognized, then retrain the model with the new face using the label number returned from the recognition
4. If not add this face to memory with the tracking number
5. Go through the list of all face images in memory to find if there are more images with the same tracking number.
6. If there are more than 5 images for the user, retrain the model with these new faces using the tracking number as the label number.

7. Remove the trained faces from memory.

The below figure shows a list of labels in the model file, this represent the labels used for each user where each number is the tracking id of a user. From this we can conclude that the model is trained on 3 users so far.


```xml
<labels type_id="opencv-matrix">
  <rows>190</rows>
  <cols>1</cols>
  <dt>i</dt>
  <data>
    7712 7712 7712 7712 7712 7712 7712 7712 7712 7730 7730 7730 7730
    7730 7712 7712 7712 7712 7712 7712 7712 7712 7712 7712 7712 7712
    7712 7712 7712 7712 7730 7730 7730 7730 7730 7712 7730 7712 7730
    7730 7730 7730 7730 7730 7712 7712 7712 7712 7712 7712 7712 7712
    7712 7712 7712 7730 7712 7730 7712 7730 7712 7730 7730 7730 7730
    7730 7730 7730 7730 7730 7730 7730 7730 7730 7730 7730 7730 7730
    7730 7730 7730 7712 7712 7712 7712 7712 7712 7712 7712 7712 7712
    7712 7712 7712 7712 7712 7712 7712 7712 7738 7738 7738 7738 7738
    7730 7712 7712 7712 7730 7712 7730 7730 7712 7712 7730 7712 7712
    7730 7730 7730 7730 7730 7730 7730 7730 7730 7730 7921 7921 7921
    7921 7921 7921 7712 7712 7712 7712 7712</data></labels>
```


Figure 27 LBPH Model

5. Proposed Methods for Object Recognition

The goal of this proposed method is to detect and identify objects in a given scene.

These scene images are static images loaded from local storage.

5.1 Loading and pre-processing known objects

The proposed method starts with loading and preparing all the known objects, which is done in preparation for the use of feature extraction and histogram comparison to be used later on.

(Coloured) Images are loaded into memory where each image has a name and a number. Images are resized to 50x150 pixels.

After all images of known objects are loaded into memory, scene images are then loaded one by one.

5.2 Pre-processing scene images

The scene is an image of a fridge shelf with drinks (cans or bottles) on it.

Pre-processing is required to remove any background in the image and also locate a ROI where it is showing only the drinks on the shelf.

This will help improve the feature extraction and histogram comparisons as most of the noise in the scene has been removed.

5.2.1 Sharpening the scene image

This is done by smoothing the image using Gaussian filter and then calculating the absolute difference of the smoothed and original images.

Images are then resized to 500x500 pixels

Figure 28 Sharpened Scene Image

5.2.2 background/noise removal

As seen in figure 28, there is a lot of noise in the scene images (bottles from the lower shelf, the back and sides of the fridge etc...).

The goal is to remove all the noise and only keep the drink objects in the scene as much as possible by applying the following proposed algorithm:

1. Find all the edges in the scene using Canny edge detector
2. Find all the lines from in the image using the Hough transform
3. Go through all the lines and discard any vertical lines and only keep the horizontal ones.
4. From the horizontal lines only keep the longest lines as they represent the shelf edge
5. Sort the horizontal lines based on their location in the scene
6. From the top half of the scene discard anything above the lowest horizontal line
7. From the bottom half of the scene, discard anything lower than the highest horizontal line.
8. Finally, cut off 10% of the image from both sides as this would show the vertical sides of the fridge

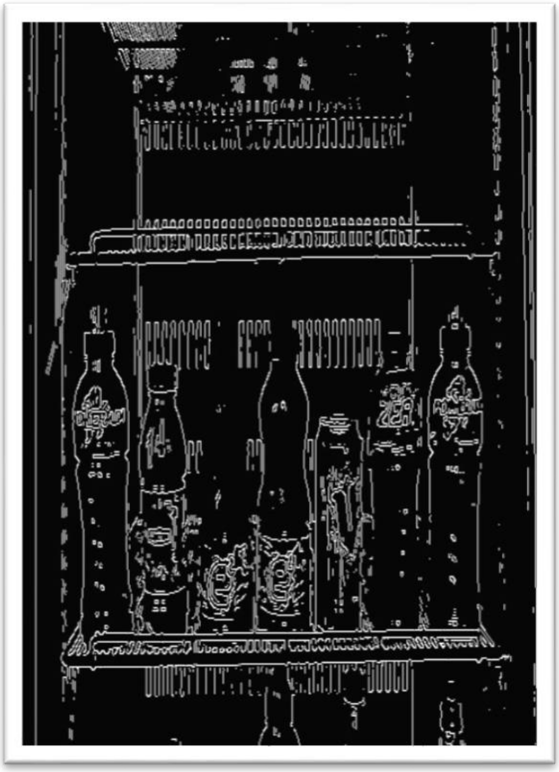


Figure 29 Canny Edge Detection

Figure 30 Hough transform to find lines

Figure 31 Longest Horizontal Lines

Figure 32 Final Pre-processed image

5.3 Object detection/localization

The goal is to detect and locate the objects in the scene, which in this case are cans and bottles on the shelf.

It is given that the shelf can hold up to 7 objects wide, based on this the scene (processed scene with no background or noise) will be divided into 7 images vertically (see table 2).

Table 2 Object Localization

For each of the divided images, the middle area will be extracted. The reason for this is to further enhance the removal of the background and noise from the objects and to only be left with an area that best describes the objects which will be used later for object classification, see table 3

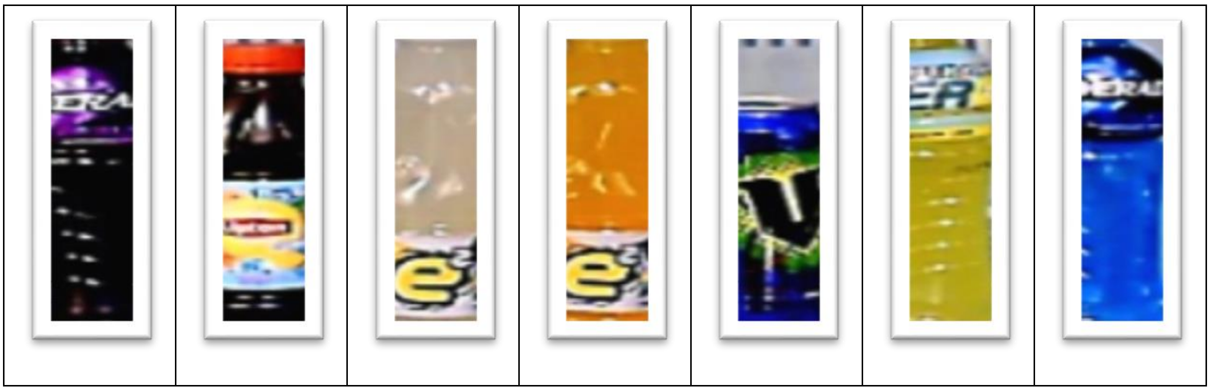


Table 3 Object Detection

5.4 Object classification

The proposed method for object classifications consists of two techniques, one is to use Histogram to compare between the known object images and the located object and find the best/top 2 matches.

The second is to further enhance the top 2 matches by using feature extraction and find the best/top match.

5.4.1 Histogram comparison

Once an object is located and detected (from table 3), the next step is to calculate the histogram of this object and compare it to the know objects in the system, below describes the steps taken to achieve this:

1. Convert the known image to HSV colour space
2. Convert the detected object image to HSV colour space
3. Calculate the histogram for the both of the above
4. Compare the two histogram and store the result in memory

Each known object in the system has multiple images (these images are taken for the object from multiple angles), therefore the above proposed algorithm is executed for these multiple images.

The one that best matches the detected object will be stored and compared to the other known objects later on to find the top 2 know objects that matched the detected object.

Histogram comparison is carried out using the Correlation method,

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H_1})(H_2(I) - \bar{H_2})}{\sqrt{\sum_I (H_1(I) - \bar{H_1})^2 \sum_I (H_2(I) - \bar{H_2})^2}}$$

Where

$$\bar{H_k} = \frac{1}{N} \sum_J H_k(J)$$

And **N** is the total number of histogram bins.

Once the best image that represents a known object in the system is chosen, the results of the Histogram comparison will be sort descending and the top 2 matches will be stored (to be processed further using feature extraction to determine the best match).

5.4.2 Feature extraction

There are 2 objects that best matches the detected object in the scene, from running experiments on this, it was concluded that most of the time the correct object will always be in those top 2 matches.

Feature extraction is used to further enhance the results of the histogram comparison and to determine the correct match as the top choice.

There are 2 methods that were investigated and used for feature extraction, these are SIFT [53] and SURF [54].

SIFT stands for Scale invariant feature transform, which extracts key points and compute its descriptors. There are four main steps that are used in SIFT:

1. Scale-space Extrema Detection

    When an image is scaled, smaller corners can be detected but to detect larger corners we will need larger windows. Scale-space is used for this. Laplacian of Gaussian is found for the image with various σ values. LoG acts as a blob detector which detects blobs in various sizes due to change in σ. Gaussian kernels with low σ give a high value for small corners while Gaussian kernel with high σ fits well for larger corners. So, we can find the local maxima across the scale and space which gives us a list of (x, y, σ) values which means there is a potential keypoint at (x, y) at σ scale.

    Gaussian Kernel

$$g_\sigma(\mathbf{x}) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{1}{2}\frac{\mathbf{x}^\top\mathbf{x}}{\sigma^2}\right)$$

The Gaussian scale space is a collection of smoothed images

$$I_\sigma = g_\sigma * I, \quad \sigma \geq 0.$$

Scale space is computed by

$$I_\sigma = g_{\sqrt{\sigma^2-\sigma_n^2}} * I_{\sigma_n}, \quad \sigma \geq \sigma_n.$$

2. Keypoint Localization

   Candidate keypoints are selected to get more accurate results, where a detailed model is matched to determine location and scale

3. Orientation assignment

   One or multiple orientations are assigned to each keypoint to achieve invariance to image rotation.

4. Keypoint descriptor

   Keypoints Descriptors are then created.

SURF stands for "Speeded up Robust Features". This algorithm is based on a faster version of SURF. Surf is faster as it approximate LoG with Box filter, one advantage of this is that convolution with box filter can be calculated with integral images (for different scales this can be done in parallel).

The following proposed algorithm details steps taken to use feature extraction techniques:

1. Features are extracted from the top 2 histogram matches and the detected object using SIFT or SURF.
    a. SIFT and SURF are being used to find out the best method to use (results are shown in Chapter 6).
2. Match the extracted feature from the 2 top object with the features extracted from the detected object
    a. Only use the strong features and ignore all the weak ones.

5.5 Summary

This chapter discussed the proposed method for object recognition and how it was implemented, where the proposed method uses a combination of two algorithms which are feature extraction and histogram. In the results and discussion chapter, the results will demonstrate how this proposed method has improved the object recognition process.

6. Results and Analysis

The experiments shown in this chapter will cover both face and object recognition.

The testing was carried on a laptop with the below specifications:

- CPU: Intel Core, I7-3630QM CPU @ 2.4 GHz
- RAM: 12GB
- Graphics Hardware: Intel HD Graphics 4000, 2GB
- OS: Windows 8, 64bit
- USB: 3.0
- Camera Specs:
    - Colour Camera: 1920 x 1080 @ 30 fs
    - Depth camera: 512 x 424
    - Max depth distance ~4.5 M
    - Min depth distance: 50 cm
    - Horizontal FOV: 70 degrees
    - Vertical FOW: 60 degrees
    - USB required: 3.0

6.1 Face Recognition Experiment

The Camera was placed about 2.2 meters above the ground and tilted down to capture the faces, Kinect 2 camera was set to only capture up to three meters away and everything beyond that was discarded.

Figure 2 Camera installed 2.2 Meter above the ground

The experiment was carried out in an environment where no control over lighting, position of people or direction of where they were looking was set.

The camera was installed for few days first to ensure participants become familiar with its location which in turn decreased the amount of time they tended to look up at the camera when they passed by. The recording was done at random times, with sufficient lighting.

Two videos were recorded with at least 2 people present in each video, where height ranged from 150 to 195 cm.

Video 1	**Not Using Skin Detection**	**Using Skin Detection**
Total Number of Frames	479	377
Number of Frames were one or more person is present	477	166
Number of frames were one or more person is detected	204	160
Number of false detections	3	0
Faces Recognized	163	136
False Recognition	15	3

Table 4 Video 1 results

<u>Video 1 Analysis (No Skin Detection):</u>

Detection recall is 42% and the precision is 98%, while the recognition task had a recall of 79% and a precision of 90%

<u>Video 1 Analysis (Using Skin Detection):</u>

Detection recall is 96% and the precision is 100%, while the recognition task had a recall of 85% and a precision of 97%

Video 2	Not Using Skin Detection	Using Skin Detection
Total Number of Frames	237	475
Number of Frames were one or more person is present	232	187
Number of frames were one or more person is detected	64	124
Number of false detections	12	14
Faces Recognized	21	71
False Recognition	2	5

Table 5 Video 2 results

Video 2 Analysis (No Skin Detection):

Detection recall is 27% and the precision is 81%, while the recognition task had a recall of 32% and a precision of 90%

Video 2 Analysis (Using Skin Detection):

Detection recall is 66% and the precision is 88%, while the recognition task had a recall of 57% and a precision of 92%

6.2 Object Recognition Experiment

The Camera was placed on the door panel where it can capture the whole shelf.

The experiments were carried out in two environments; one with no control over lighting and another where the lighting was set to make sure that it does not have any negative effect on the images. The detailed results are shown in the appendix.

6.2.1 Using SIFT and No lighting control present

Image Name	Positive Matches	False Positive Matches
1a	3	4
1b	5	2
1c	5	2
1d	4	3
1e	4	3
2a	2	5
2b	2	5
2c	1	6
2d	1	6
2e	2	5
3a	3	4
3b	2	5

3c	2	5
3d	2	5
3e	0	7
4a	2	5
4b	3	4
4c	2	5
4d	1	6
4e	2	5
5a	2	5
5b	4	3
5c	3	4
5d	4	3
5e	4	3

Table 6 Sift and no light control results

Total number of objects is: 175

Total number of correct objects recognition is: 65

Total number of wrong objects recognition is: 110

Success percentage is 37%

6.2.2 Using SIFT and lighting control present

Image Name	Positive Matches	False Positive Matches
1a	5	2
1b	5	2
1c	4	3
1d	4	3
1e	6	1
2a	3	4
2b	4	3
2c	5	2
2d	4	3
2e	4	3
3a	2	5
3b	1	6
3c	3	4
3d	2	5
3e	1	6
4a	4	3
4b	4	3
4c	4	3
4d	4	3
4e	3	4

5a	4	3
5b	6	1
5c	5	2
5d	5	2
5e	6	1

Table 7 Sift and light control present Results

Total number of objects is: 175

Total number of correct objects recognition is: 98

Total number of wrong objects recognition is: 77

Success percentage is 56%

6.2.3 Using SURF and No lighting control present

Image Name	Positive Matches	False Positive Matches
1a	5	2
1b	5	2
1c	5	2
1d	3	4
1e	4	3
2a	2	5
2b	2	5
2c	1	6
2d	1	6
2e	4	3
3a	3	4
3b	2	5
3c	1	6
3d	2	5
3e	1	6
4a	2	5
4b	2	5
4c	2	5
4d	1	6
4e	3	4

5a	1	6
5b	4	3
5c	3	4
5d	4	3
5e	3	4

Table 8 SURF and no light control Results

Total number of objects is: 175

Total number of correct objects recognition is: 66

Total number of wrong objects recognition is: 109

Success percentage is 37%

6.2.4 Using SURF and lighting control present

Image Name	Positive Matches	False Positive Matches
1a	5	2
1b	5	2
1c	4	3
1d	5	2
1e	6	1
2a	3	4
2b	5	2
2c	5	2
2d	4	3
2e	4	3
3a	2	5
3b	2	5
3c	3	4
3d	1	6
3e	2	5
4a	4	3
4b	4	3
4c	4	3
4d	4	3
4e	3	4

5a	5	2
5b	6	1
5c	5	2
5d	5	2
5e	6	1

Table 9 SURF and Lighting control present Results

Total number of objects is: 175

Total number of correct objects recognition is: 102

Total number of wrong objects recognition is: 73

Success percentage is 58%

6.2.5 Results Summary

These results will be discussed and compared further in the next chapter.

	Lighting control present	No Lighting Control
Surf	58%	37%
Sift	56%	37%

7. Conclusion and Discussion

7.1 Face Recognition

Video 1 analysis shows that using the skin detection algorithm results in a better detection recall 96% comparing to 42%. The detection precision is very similar at about ~98%-100%.

When looking at the number of frames captured, the skin detection algorithm will have less frame captured and this is due to the extra computational load using more resources and so reducing the amount of frames that can be captured.

There is a slight improvement in the recognition process, when using the skin detection algorithm it results in an 85% recall and 97% precision comparing to 79% and 90% precision when using the non-skin detection algorithm. The skin detection algorithm provided better segmented face images to the recognition process as a lot more of the background noise is removed.

Video 2 shows a significant improvement in the detection process using the skin detection algorithm, where the detection recall is 66% comparing to 27% and precision is 88%. This improved recall and precision also resulted in a better recognition recall with 57% when using the skin detection algorithm comparing to 32% and a precision of 92%.

The proposed method using skin detection and depth information (where the users location is identified and used for tracking and correcting any false positive results from the skin detection algorithm) shows significant improvement in face recognition, especially in environments when there is no control over the lighting and were people are walking in front of the camera and not facing it in some cases.

7.2 Object Recognition

The experiment compared the proposed method when carried out in a controlled lighting environment to a non-controlled lighting environment, it also compared the performance of the SURF and SIFT feature extraction algorithms.

As expected in the controlled lighting environment, the proposed approach shows better results with 58% comparing to 37% while SURF and SIFT produced equal results.

From a performance point of view, it would be better to use the SURF algorithm as it is faster to extract features for the same accuracy.

The difference in results between controlled and un-controlled lighting environment is acceptable in the context of this research.

Following is a comparison of the results from this research to results carried out by other research in similar areas. The ILSVRC-2010 competition best performance results was 47.1& and 28.2% [55] and since then the best published results are 45.7 and 25.7 [56]. These experiments were carried out on datasets where the images were of good quality and varying lighting conditions were not present as in this research.

7.3 Future Work

We will investigate improving the results of face recognition under even more widely varying illumination levels using a new state of the art camera such as the new Intel RealSense camera (R200) which supports depth data up to 10 meters in sunlight and dimly lit indoor lighting conditions. This is a reasonably priced camera for ~$99 USD and would be robust to sun light streaming into a retail environment.

We will extend the the Object classification and compare our proposed method with CNN (deep learning) using the torch framework.

References

[1] G. Eason, B. Noble, I.N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions”, Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955.

[2] J. Clerk Maxwell, “A Treatise on Electricity and Magnetism”, 3rd ed., vol. 2, Oxford: Clarendon, pp. 68-73, 1892.

[3] I.S. Jacobs, C.P. Bean, “Fine particles, thin films and exchange anisotropy”, in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, pp. 271-350, 1963.

[4] J. N. K. Liu, M. Wang, B. Feng, "iBotGuard: an Internet-based intelligent robot security system using invariant face recognition against intruder", IEEE Transactions on Systems Man and Cybernetics PartC-Applications and Reviews, Vol.35, pp.97-105, 2005.

[5] T. Jebara, "3D Pose Estimation and Normalization for Face Recognition", Center for Intelligent Machines, McGill University, May, 1996

[6] M. Yang, D.J. Kriegman, N. Ahuja, “Detecting Faces in Images: A Survey”, IEEE Transactions on a pattern analysis and machine intelligence, Vol. 24 No. 1, January 2002.

[7] M. Modi, F. Macwan, “Face Detection Approaches: A Survey”, International Journal of Innovation Research in Science, Engineering and technology, Vol. 3, Issue 4, April 2014.

[8] G. Yang,T. S. Huang, “Human Face Detection in Complex Background”, Pattern Recognition, vol. 27, no. 1, pp. 53-63, 1994.

[9] C. Kotropoulos, I. Pitas, “Rule-Based Face Detection in Frontal Views”, Proc. Int’l Conf. Acoustics, Speech and Signal Processing, vol. 4, pp. 2537-2540, 1997.

[10] S.A. Sirohey, “Human Face Segmentation and Identification”, Technical Report CS-TR-3176, Univ. of Maryland, 1993.

[11] H.P. Graf, T. Chen, E. Petajan, E. Cosatto, “Locating Faces and Facial Parts”, Proc. First Int’l Workshop Automatic Face and Gesture Recognition, pp. 41-46, 1995.

[12] M.F. Augusteijn, T.L. Skujca, “Identification of Human Faces through Texture-Based Feature Recognition and Neural Network Technology”, Proc. IEEE Conf. Neural Networks, pp. 392-398, 1993.

[13] H.P. Graf, T. Chen, E. Petajan, E. Cosatto, “Locating Faces and Facial Parts”, Proc. First Int’l Workshop Automatic Face and Gesture Recognition, pp. 41-46, 1995.

[14] H.P. Graf, E. Cosatto, D. Gibbon, M. Kocheisen, E. Petajan, “Multimodal System for Locating Heads and Faces”, Proc. Second Int’l Conf. Automatic Face and Gesture Recognition, pp. 88-93, 1996.

[15] J. Yang, A. Waibel, “A Real-Time Face Tracker”, Proc. Third Workshop Applications of Computer Vision, pp. 142-147, 1996.

[16] Z. Jin, Z. Lou, J. Yang, Q. Sun, “Face detection using template matching and skin-color information”, Neurocomputing, vol. 70, no. 4-6, pp. 794-800, 2007.

[17] P. Kakumanu, S. Makrogiannis, N. Bourbakis, “A survey of skin-color modeling and detection methods”, Pattern Recognition, vol. 40, no. 3, pp. 1106-1122, 2007.

[18] A.B. Usakli, A. Susac, S. Gurkan, “Fast face recognition: Eye blink as a reliable behavioral response”, Neuroscience Letters, vol. 504, no. 1, pp. 49-52, 2011.

[19] C. Kervrann, F. Davoine, P. Perez, H. Li, R. Forchheimer, C. Labit, “Generalized Likelihood Ratio-Based Face Detection and Extraction of Mouth Features”, Proc. First Int’l Conf. Audio- and Video-Based Biometric Person Authentication, pp. 27-34, 1997

[20] M.-H. Yang, N. Ahuja, D. Kriegman, "Mixtures of Linear Subspaces for Face Detection", Proc. Fourth Int'l Conf. Automatic Face and Gesture recognition, pp. 70-76, 2000.

[21] A. J. Colmenarez, T. S. Huang, "Face detection with information-based maximum discrimination", in IEEE Proc. of Int. Conf. on CVPR, 1997.

[22] E. Osuna, R. Freund, F. Girosi, "Training Support Vector Machines: An Application to Face Detection", Proc. IEEE Conf. CVPR, pp. 130-136, 1997

[23] H. Schneiderman, T. Kanade, "Probabilistic Modeling of Local Appearance and Spatial Relationships for Object Recognition", Proc. IEEE Conf. CVPR, pp. 45-51, 1998

[24] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", Proc. IEEE Conference on CVPR, vol. 1, pp. 511-518, 2001.

[25] P. I. Wilson, J. Fernandez, "Facial feature detection using Haar classifers", Journal of Computing Sciences in Colleges, vol. 21, no. 4, pp. 127-133, 2006.

[26] M. Turk,A. Pentland, "Eigenfaces for recognition", Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[27] Kyungnam Kim, “Face recognition using Principle Component Analysis.” Computer Vision Lab, University of Maryland, 2012

[28] Aleix M. Martinez, Avinash C. Kak, “PCA versus LDA”, IEEE Trans. Pattren Analysis and Machine Intelligence, Vol. 23, No. 2, pp.228-233, 2001.

[29] X. Wang, C. Huang, X. Fang, J. Liu, ” 2DPCA vs. 2DLDA: Face Recognition using Two-Dimensional Method”, International Conference on Artificial Intelligence and Computational Intelligence, 2009.

[30] J. Yang , D. Zhang, A. F. Frangi, J.Y. Yang, “Two- dimensional PCA: a new approach to appearance-based face representation and recognition”, IEEE Trans. On Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 131-137, Jan. 2004.

[31] Wiskott, L., Fellous, J., Krüger, N., Malsburg, “Face Recognition By Elastic Bunch Graph Matching”, IEEE Transactions on Pattern Analysis and Machine Intelligence 19, S. 775—779, 1997.

[32] Messer, K. et al., “Performance Characterisation of Face Recognition Algorithms and Their Sensitivity to Severe Illumination Changes”, ICB, S. 1—11, 2006.

[33] T. Ahonen , A. Hadid, M. Pietikainen, "Face Recognition with Local Binary Patterns", Computer Vision - ECCV, 469—481, 2004.

[34] Y. Rodriguez, "Face detection and verification using Local Binary Patterns", Ecole Polytechnique Federale de Lausanne, 2006.

[35] T.Ahonen, A. Hadid, M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition", IEEE Transactions on a pattern analysis and machine intelligence. Vol. 28, issue 12, 2006.

[36] H. Schneidernam, T. Kanade; "Object Detection Using the Statistics of Parts"; International Journal of Computer Vision, pp 151-177, 2004.

[37] C. H. Lampert, M. B. Blaschko, T. Hofmann, "Beyond sliding windows:Object localization by efficient subwindow search", CVPR 2008, pp. 1—8, 2008.

[38] Li He, Hui Wang, Hong Zhang, "Object detection by parts using appearance, Structural and shape features", IEEE Conference Mechatronics and Automation(ICMA), 2011; page(s): 489 - 494 ; 2011.

[39] M. Weber, et al. “Towards automatic discovery of object categories”, CVPR, 2000.

[40] N. Dalal, B. Triggs, “Histograms of oriented gradients for human detection”, CVPR 2005

[41] H. S. Parekh, D. G. Thakore, U. K. Jaliya; “A Survey on Object Detection and Tracking Methods”, International Journal of Innovation Research in Computer and Communication Engineering, Vol.2, Issue 2, Feb 2014

[42] S. S. M. Radzi, S. N. Yaakob , Z. Kadim, H. H. Woon, “Extraction of Moving Objects Using Frame Differencing, Ghost and Shadow Removal”, 5th International Conference on Intelligent Systems, Modelling and Simulation, pp. 229-234, 2014

[43] A. K. Chauhan, P. Krishan, “Moving Object Tracking Using Gaussian Mixture Model And Optical Flow”, International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, issue 4, April 2013.

[44] M. Swain, D. Ballard, “Color indexing,” International journal of computer vision, 1991.

[45] D. G. Lowe, "Object recognition from local scale-invariant features", proceedings of the seventh IEEE international conference on Computer vision, vol. 2, pp. 1150—1157, 1999.

[46] H.Yue, X. Sun, F. Wu, J. Yang "SIFT- Based Image Compression" IEEE ICME, pp 473-478, July, 2012

[47] A. Alfanindya, N. Hashim, C. Eswaran; "A Survey on Object Detection and Tracking Methods" International Conference on technology, Informatics, Management, Engineering & Environment, June, 2013.

[48] http://docs.opencv.org/2.4/doc/tutorials/features2d/feature_description/feature_description.html, Nov, 2015

[49] Marius Muja, David G. Lowe. "Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration", 2009

[50] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Recognition with Local Binary Patterns", Computer Vision - ECCV, 469—481, 2004.

[51] T.Ahonen, A. Hadid, M. Pietikainen ," Face Description with Local Binary Patterns: Application to Face Recognition", IEEE Transactions on a pattern analysis and machine intelligence. Vol. 28, issue 12, 2006.

[52] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, D. Ramanan, "Object Detection with Discriminatively Trained Part-Based Models", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, Issue. 9, 2010

[53] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoint", International Journal of Computer Vision, 2004.

[54] H. Baya, A. Essa, T. Tuytelaarsb, L. Goola, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding, Vol. 110, Issue 3, pp. 346-359. 2007

[55] A. Berg, J. Deng, and L. Fei-Fei. Large scale visual recognition challenge 2010. www.imagenet.org/challenges. 2010

[56] J. Sánchez and F. Perronnin. High-dimensional signature compression for large-scale image classification. In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pages 1665—1672. IEEE, 2011.

Appendix

Using SIFT and No lighting control present

Image Name	Object Number	Histogram Top 2 results	Feature Extraction Result	Positive (P) /False Positive (FP)	The correct Result is
1a	1	PowerRadePurple Lift Green	Lift Green	FP	PowerRadePurple
	2	PowerRadeRed Lipton	PowerRadeRed	FP	Lipton
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	Lipton RedbullSugar	Lipton	FP	V Blue
	6	Lift GForce	Lift	FP	PowerRade Zero
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
1b	1	LiftGreen PowerRadePurple	LiftGreen	FP	PowerRadePurple
	2	Lipton LiftGreen	Lipton	P	
	3	E2Lemon PowerRadeZero	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	Lift PowerRadeZero	Lift	FP	PowerRadeZero
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
1c	1	VBlue LiftGreen	VBlue	FP	PowerRadePurple
	2	Lipton PowerRadeRed	Lipton	P	
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	Lift PowerRadeZero	Lift	FP	PowerRadeZero
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
1d	1	LiftGreen CokeZero	CokeZero	FP	PowerRadePurple
	2	Lipton RedbullSugar	Lipton	P	
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	Lift	GForce	FP	PowerRadeZero

		GForce			
	7	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
1e	1	CokeZero LiftGreen	CokeZero	FP	PowerRadePurple
	2	Lipton PowerRadeRed	Lipton	P	
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	Redbull VBlue	Redbull	FP	VBlue
	6	Lift GForce	Lift	FP	PowerRadeZero
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
2a	1	PowerRadeRed Mother	PowerRadeRed	FP	Coke
	2	V CokeZero	V	P	
	3	E2Lemon Mother	E2Lemon	FP	Mother
	4	PowerRadeZero E2Lemon	PowerRadeZero	FP	Lift
	5	Mother RedbullSugar	RedbullSugar	FP	CokeZero
	6	VBlue Redbull	VBlue	FP	RedBull
	7	GForce V	GForce	P	
2b	1	LiftGreen Redbull	LiftGreen	FP	Coke
	2	V Lift	V	P	
	3	Mother PowerRadeRed	Mother	P	
	4	PowerRadeZero E2Orange	PowerRadeZero	FP	Lift
	5	LiftGreen Mother	LiftGreen	FP	CokeZero
	6	Redbull RedbullSugar	RedbullSugar	FP	RedBull
	7	V GForce	V	FP	GForce
2c	1	PowerRadeRed Mother	PowerRadeRed	FP	Coke
	2	GForce V	GForce	FP	V
	3	E2Lemon Mother	E2Lemon	FP	Mother
	4	E2Orange Hopt	E2Orange	FP	Lift
	5	E2Lemon Redbull	Redbull	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	FP	RedBull
	7	GForce V	GForce	P	
2d	1	Lipton LiftGreen	Lipton	FP	Coke
	2	V Lift	V	P	

	3	LiftGreen Lipton	LiftGreen	FP	Mother
	4	PowerRadeZero E2Lemon	PowerRadeZero	FP	Lift
	5	LiftGreen Mother	LiftGreen	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	FP	RedBull
	7	V Lift	v	FP	GForce
2e	1	Lipton PowerRadeRed	Lipton	FP	Coke
	2	V GForce	V	P	
	3	Coke PowerRadePurple	Coke	FP	Mother
	4	E2Orange PowerRadeZero	E2Orange	FP	Lift
	5	LiftGreen Mother	LiftGreen	FP	CokeZero
	6	Redbull VBlue	VBlue	FP	RedBull
	7	GForce V	GForce	P	
3a	1	LiftGreen Redbull	LiftGreen	FP	Hop
	2	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
	3	Mother CokeZero	CokeZero	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	CokeZero LiftGreen	CokeZero	P	
	6	Lipton Redbull	Lipton	FP	RedBullSugar
	7	Redbull LiftGreen	Redbull	FP	Water
3b	1	CokeZero Mother	CokeZero	FP	Hopt
	2	VBlue RedbullSugar	VBlue	FP	PowerRadeBlue
	3	Mother E2Lemon	Mother	P	
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	Redbull Lipton	Redbull	FP	RedBullSugar
	7	Redbull RedbullSugar	Redbull	FP	Water
3c	1	LiftGreen Redbull	LiftGreen	FP	Hopt
	2	VBlue Redbull	VBlue	FP	PowerRadeBlue
	3	LiftGreen Mother	LiftGreen	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	CokeZero Mother	CokeZero	P	
	6	Redbull	Redbull	FP	RedBullSugar

		Lipton			
	7	RedbullSugar Redbull	RedbullSugar	FP	Water
3d	1	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	2	VBlue RedbullSugar	VBlue	FP	PowerRadeBlue
	3	E2Lemon Mother	Mother	P	
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen Redbull	LiftGreen	FP	CokeZero
	6	Lipton Redbull	Lipton	FP	RedBullSugar
	7	Redbull RedbullSugar	Redbull	FP	Water
3e	1	Lipton Mother	Lipton	FP	Hopt
	2	VBlue RedbullSugar	VBlue	FP	PowerRadeBlue
	3	CokeZero Mother	CokeZero	FP	Mother
	4	PowerRadeRed Lipton	Lipton	FP	PowerRadeRed
	5	LiftGreen Redbull	Redbull	FP	CokeZero
	6	Lipton Redbull	Redbull	FP	RedBullSugar
	7	Redbull RedbullSugar	Redbull	FP	Water
4a	1	E2Lemon E2Orange	E2Lemon	P	
	2	CokeZero Mother	CokeZero	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	E2Lemon Mother	E2Lemon	FP	Mother
	5	CokeZero Mother	CokeZero	FP	Hopt
	6	Redbull Lipton	Redbull	FP	LiftGreen
	7	Redbull RedbullSugar	Redbull	FP	Lipton
4b	1	E2Lemon E2Orange	E2Lemon	P	
	2	PowerRadeRed Lipton	PowerRadeRed	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Lipton PowerRadeRed	Lipton	FP	Mother
	5	PowerRadeRed Lipton	PowerRadeRed	FP	Hopt
	6	Lipton Redbull	Redbull	FP	LiftGreen
	7	RedbullSugar Lipton	Lipton	P	
4c	1	E2Lemon E2Orange	E2Lemon	P	
	2	CokeZero Lipton	CokeZero	FP	Water

	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	LiftGreen Lipton	Lipton	FP	Mother
	5	CokeZero Moth	CokeZero	FP	Hopt
	6	Lipton Redbull	Redbull	FP	LiftGreen
	7	RedbullSugar Lipton	RedbullSugar	FP	Lipton
4d	1	E2Lemon E2Orange	E2Lemon	P	
	2	E2Lemon LiftGreen	E2Lemon	FP	Water
	3	Lift PowerRadeZero	Lift	FP	PowerRadeZero
	4	Coke Hopt	Hopt	FP	Mother
	5	Lipton PowerRadeRed	Lipton	FP	Hopt
	6	Redbull Lipton	Redbull	FP	LiftGreen
	7	Redbull RedbullSugar	Redbull	FP	Lipton
4e	1	E2Lemon E2Orange	E2Lemon	P	
	2	LiftGreen Mother	LiftGreen	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Coke Mother	Coke	FP	Mother
	5	PowerRadeRed Lipton	Lipton	FP	Hopt
	6	Lipton Redbull	Redbull	FP	LiftGreen
	7	Redbull LiftGreen	Redbull	FP	Lipton
5a	1	Hopt E2Orange	Hopt	FP	E2Orange
	2	LiftGreen Lipton	LiftGreen	FP	PowerRadePurple
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V Lift	V	P	
	5	LiftGreen Lipton	Lipton	FP	Hopt
	6	PowerRadeRed Lipton	PowerRadeRed	FP	Coke
	7	Redbull RedbullSugar	Redbull	FP	Lipton
5b	1	E2Orange Hopt	E2Orange	P	
	2	LiftGreen Water	LiftGreen	FP	PowerRadePurple
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V Lift	V	P	
	5	PowerRadeRed Lipton	Lipton	FP	Hopt
	6	Mother	CokeZero	FP	Coke

		CokeZero			
	7	Lipton RedbullSugar	Lipton	P	
5c	1	E2Orange Hopt	E2Orange	P	
	2	LiftGreen PowerRadePurple	LiftGreen	FP	PowerRadePurple
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V Lift	V	P	
	5	PowerRadeRed Lipton	Lipton	FP	Hopt
	6	VBlue Redbull	VBlue	FP	Coke
	7	LiftGreen Redbull	Redbull	FP	Lipton
5d	1	E2Orange Hopt	E2Orange	P	
	2	LiftGreen PowerRadePurple	LiftGreen	FP	PowerRadePurple
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V CokeZero	V	P	
	5	Hopt PowerRadeRed	Hopt	P	
	6	Redbull VBlue	Redbull	FP	Coke
	7	Redbull RedbullSugar	Redbull	FP	Lipton
5e	1	E2Orange Hopt	E2Orange	P	
	2	LiftGreen Lipton	LiftGreen	FP	PowerRadePurple
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V LiftGreen	V	P	
	5	Hopt CokeZero	Hopt	P	
	6	LiftGreen Mother	LiftGreen	FP	Coke
	7	LiftGreen Mother	LiftGreen	FP	Lipton

Using SIFT in a controlled lighting environment

Image Name	Object Number	Histogram Top 2 results	Feature Extraction Result	Positive (P) /False Positive (FP)	The correct Result is
1a	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Coke Mother	Coke	FP	Lipton
	3	E2Lemon Mother	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	

	5	VBlue Redbull	VBlue	P	
	6	PowerRadeZero Lift	PowerRadeZero	P	
	7	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
1b	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Mother Water	Mother	FP	Lipton
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	PowerRadeZero Lift	PowerRadeZero	P	
	7	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
1c	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Water Coke	Coke	FP	Lipton
	3	E2Lemon PowerRadeZero	E2Lemon	P	
	4	E2Orange E2Lemon	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	Lift PowerRadeZero	Lift	FP	PowerRadeZero
	7	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
1d	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Water Coke	Water	FP	Lipton
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange E2Lemon	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	Lift PowerRadeZero	Lift	FP	PowerRadeZero
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	FP	PowerRadeBlue
1e	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Water Mother	Water	FP	Lipton
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	PowerRadeZero Lift	PowerRadeZero	P	
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	

2a	1	Coke Water	Coke	P	
	2	CokeZero V	CokeZero	FP	V
	3	CokeZero Mother	CokeZero	FP	Mother
	4	Lift PowerRadeZero	Lift	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	VBlue Redbull	VBlue	FP	RedBull
	7	GForce V	GForce	P	
2b	1	Coke Water	Coke	P	
	2	V GForce	V	P	
	3	Water Coke	Water	FP	Mother
	4	Lift PowerRadeZero	Lift	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	Redbull LiftGreen	LiftGreen	FP	RedBull
	7	GForce V	GForce	P	
2c	1	Coke Water	Coke	P	
	2	V GForce	V	P	
	3	CokeZero Mother	CokeZero	FP	Mother
	4	Lift PowerRadeZero	Lift	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	Redbull LiftGreen	Redbull	P	
	7	GForce V	GForce	P	
2d	1	Coke Water	Coke	P	
	2	V LiftGreen	V	P	
	3	Water Coke	Water	FP	Mother
	4	Lift PowerRadeZero	Lift	P	
	5	LiftGreen CokeZero	LiftGreen	P	
	6	Redbull LiftGreen	Redbull	FP	CokeZero
	7	GForce V	GForce	FP	RedBull
2e	1	Coke Water	Coke	P	
	2	V GForce	V	P	
	3	CokeZero Mother	CokeZero	FP	Mother
	4	Lift	Lift	P	

		PowerRadeZero			
	5	LiftGreen CokeZero	LiftGreen	P	
	6	VBlue Redbull	VBlue	FP	CokeZero
	7	GForce V	GForce	FP	RedBull
3a	1	Mother Water	Mother	FP	Hop
	2	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
	3	Water Mother	Water	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	RedbullSugar LiftGreen	RedbullSugar	P	
	7	Mother LiftGreen	Mother	FP	Water
3b	1	Coke Water	Coke	FP	Hopt
	2	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
	3	Water Coke	Water	FP	Mother
	4	PowerRadeRed Lipton	Lipton	FP	PowerRadeRed
	5	LiftGreen Mother	LiftGreen	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	FP	RedBullSugar
	7	Mother PowerRadeRed	Mother	FP	Water
3c	1	Mother Water	Mother	FP	Hopt
	2	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
	3	Mother Water	Mother	P	
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	CokeZero LiftGreen	CokeZero	P	
	6	LiftGreen CokeZero	LiftGreen	FP	RedBullSugar
	7	Mother CokeZero	Mother	FP	Water
3d	1	Mother Coke	Mother	FP	Water
	2	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
	3	Coke Water	Coke	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	P	
	7	LiftGreen RedbullSugar	LiftGreen	FP	Water

3e	1	Water Mother	Water	FP	Hopt
	2	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
	3	Water Coke	Water	FP	Mother
	4	PowerRadeRed Lipton	Lipton	FP	PowerRadeRed
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	FP	RedBullSugar
	7	LiftGreen RedbullSugar	LiftGreen	FP	Water
4a	1	E2Lemon RedbullSugar	E2Lemon	P	
	2	LiftGreen Mother	LiftGreen	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Mother Water	Mother	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	LiftGreen RedbullSugar	LiftGreen	P	
	7	Mother Coke	Mother	FP	Lipton
4b	1	E2Lemon Lift	E2Lemon	P	
	2	Mother PowerRadeRed	Mother	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Coke Water	Coke	FP	Mother
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hop
	6	LiftGreen Mother	LiftGreen	P	
	7	Lipton RedbullSugar	Lipton	P	
4c	1	E2Lemon Lift	E2Lemon	P	
	2	LiftGreen Mother	LiftGreen	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Mother Water	Mother	P	
	5	Mother PowerRadeRed	Mother	FP	Hopt
	6	LiftGreen CokeZero	LiftGreen	P	
	7	Mother Coke	Mother	FP	Lipton
4d	1	E2Lemon E2Orange	E2Lemon	P	
	2	Redbull RedbullSugar	Redbull	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Mother	Mother	P	

		Water			
	5	Mother PowerRadeRed	Mother	FP	Hopt
	6	LiftGreen RedbullSugar	LiftGreen	P	
	7	Mother Coke	Mother	FP	Lipton
4e	1	E2Orange E2Lemon	E2Orange	FP	E2Lemon
	2	LiftGreen Redbull	LiftGreen	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Coke Water	Coke	FP	Mother
	5	Mother PowerRadeRed	Mother	FP	Hopt
	6	LiftGreen CokeZero	LiftGreen	P	
	7	Lipton Water	Lipton	P	
5a	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Water	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V CokeZero	V	FP	V
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke Water	Coke	P	
	7	PowerRadeRed Mother	PowerRadeRed	FP	Lipton
5b	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Coke	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V GForce	V	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke Water	Coke	P	
	7	Lipton PowerRadeRed	Lipton	P	
5c	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Water	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V GForce	V	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke PowerRadePurple	Coke	P	
	7	Water PowerRadeRed	PowerRadeRed	FP	Lipton

5d	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Water	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V GForce	V	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke Mother	Coke	P	
	7	PowerRadeRed Mother	PowerRadeRed	FP	Lipton
5e	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Water	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V GForce	V	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke Water	Coke	P	
	7	Lipton PowerRadeRed	Lipton	P	

Using SURF and No lighting control present

Image Name	Object Number	Histogram Top 2 results	Feature Extraction Result	Positive (P) /False Positive (FP)	The correct Result is
1a	1	PowerRadePurple LiftGreen	PowerRadePurple	P	
	2	PowerRadeRed Lipton	Lipton	P	
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	Lipton RedbullSugar	Lipton	FP	VBlue
	6	Lift GForce	Lift	FP	PowerRadeZero
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
1b	1	LiftGreen PowerRadePurple	LiftGreen	FP	PowerRadePurple
	2	Lipton RedbullSugar	Lipton	P	
	3	E2Lemon PowerRadeZero	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	Lift	Lift	FP	PowerRadeZero

		PowerRadeZero			
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
1c	1	VBlue LiftGreen	VBlue		PowerRadePurple
	2	Lipton PowerRadeRed	Lipton	P	
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	Lift PowerRadeZero	Lift	FP	PowerRadeZero
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
1d	1	LiftGreen Mother	LiftGreen	FP	PowerRadePurple
	2	Lipton RedbullSugar	Lipton	P	
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	PowerRadeBlue RedbullSugar	PowerRadeBlue	FP	VBlue
	6	Lift GForce	Lift	FP	PowerRadeZero
	7	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
1e	1	LiftGreen Mother	LiftGreen	FP	PowerRadePurple
	2	Lipton Mother	Lipton	P	
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	Redbull VBlue	Redbull	FP	VBlue
	6	Lift GForce	Lift	FP	PowerRadeZero
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
2a	1	PowerRadeRed Mother	PowerRadeRed	FP	Coke
	2	V CokeZero	V	P	
	3	E2Lemon Mother	E2Lemon	FP	Mother
	4	PowerRadeZero E2Lemon	PowerRadeZero	FP	Lift
	5	Mother CokeZero	Mother	FP	CokeZero
	6	VBlue Redbull	VBlue	FP	RedBull
	7	GForce V	GForce	P	
2b	1	Mother Lipton		FP	Coke
	2	V	V	P	

		CokeZero			
	3	PowerRadeRed Mother	PowerRadeRed	FP	Mother
	4	E2Orange PowerRadeZero	E2Orange	FP	Lift
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	Redbull Mother	Redbull	P	
	7	V GForce	V	FP	GForce
2c	1	LiftGreen PowerRadeRed	LiftGreen	FP	Coke
	2	GForce V	GForce	FP	V
	3	E2Lemon Mother	E2Lemon	FP	Mother
	4	E2Orange Hopt	E2Orange	FP	Lift
	5	Redbull RedbullSugar	Redbull	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	FP	RedBull
	7	GForce V	GForce	P	
2d	1	LiftGreen Mother	LiftGreen	FP	Coke
	2	V CokeZero	V	P	
	3	LiftGreen PowerRadeRed	LiftGreen	FP	Mother
	4	PowerRadeZero E2Lemon	PowerRadeZero	FP	Lift
	5	Mother LiftGreen	Mother	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	FP	RedBull
	7	V Lift	V	FP	GForce
2e	1	Lipton PowerRadeRed	Lipton	FP	Coke
	2	V Lift	V	P	
	3	Coke PowerRadePurple	Coke	FP	Mother
	4	E2Orange PowerRadeZero	E2Orange	FP	Lift
	5	CokeZero LiftGreen	CokeZero	P	
	6	Redbull VBlue	Redbull	P	
	7	GForce V	GForce	P	
3a	1	LiftGreen CokeZero	LiftGreen	FP	Hopt
	2	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
	3	Mother Coke	Mother	P	
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero

	6	Redbull Lipton	Redbull	FP	RedBullSugar
	7	LiftGreen Redbull	LiftGreen	FP	Water
3b	1	Mother Lipton	Mother	FP	Hopt
	2	VBlue RedbullSugar	VBlue	FP	PowerRadeBlue
	3	Mother E2Lemon	Mother	P	
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	Lipton Redbull	Lipton	FP	RedBullSugar
	7	Redbull LiftGreen	Redbull	FP	Water
3c	1	LiftGreen Redbull	LiftGreen	FP	Hopt
	2	VBlue Redbull	VBlue	FP	PowerRadeBlue
	3	LiftGreen CokeZero	LiftGreen	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	Redbull Lipton	Redbull	FP	RedBullSugar
	7	RedbullSugar Lipton	RedbullSugar	FP	Water
3d	1	PowerRadeRed Lipton	PowerRadeRed	FP	Hopt
	2	VBlue PowerRadeBlue	VBlue	FP	PowerRadeBlue
	3	E2Lemon Mother	E2Lemon	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	CokeZero LiftGreen	CokeZero	P	
	6	Redbull Lipton	Redbull	FP	RedBullSugar
	7	Redbull RedbullSugar	Redbull	FP	Water
3e	1	LiftGreen CokeZero	LiftGreen	FP	Hopt
	2	VBlue Redbull	VBlue	FP	PowerRadeBlue
	3	LiftGreen Mother	LiftGreen	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen Redbull	LiftGreen	FP	CokeZero
	6	Redbull Lipton	Redbull	FP	RedBullSugar
	7	Redbull Lipton	Redbull	FP	Water
4a	1	E2Lemon E2Orange	E2Lemon	P	
	2	Mother	Mother	FP	Water

		CokeZero			
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	E2Lemon Mother	E2Lemon	FP	Mother
	5	CokeZero Mother	CokeZero	FP	Hopt
	6	Redbull Lipton	Redbull	FP	LiftGreen
	7	Redbull RedbullSugar	Redbull	FP	Lipton
4b	1	E2Lemon E2Orange	E2Lemon	P	
	2	PowerRadeRed Lipton	PowerRadeRed	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Lipton PowerRadeRed	Lipton	FP	Mother
	5	CokeZero Mother	CokeZero	FP	Hopt
	6	Lipton Redbull	Lipton	FP	LiftGreen
	7	RedbullSugar Lipton	RedbullSugar	FP	Lipton
4c	1	E2Lemon E2Orange	E2Lemon	P	
	2	Mother E2Lemon	Mother	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	LiftGreen PowerRadeRed	LiftGreen	FP	Mother
	5	CokeZero LiftGreen	CokeZero	FP	Hopt
	6	Redbull Lipton	Redbull	FP	LiftGreen
	7	RedbullSugar Lipton	RedbullSugar	FP	Lipton
4d	1	E2Lemon E2Orange	E2Lemon	P	
	2	E2Lemon LiftGreen	E2Lemon	FP	Water
	3	Lift PowerRadeZero	Lift	FP	PowerRadeZero
	4	LiftGreen Lipton	LiftGreen	FP	Mother
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Redbull Lipton	Redbull	FP	LiftGreen
	7	Redbull RedbullSugar	Redbull	FP	Lipton
4e	1	E2Lemon E2Orange	E2Lemon	P	
	2	Mother CokeZero	Mother	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	PowerRadeRed Mother	PowerRadeRed	FP	Mother
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt

	6	Lipton Redbull	Lipton	FP	LiftGreen
	7	Lipton Redbull	Lipton	P	
5a	1	Hopt E2Orange	Hopt	FP	E2Orange
	2	LiftGreen Water	LiftGreen	FP	PowerRadePurple
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Lift V	Lift	FP	V
	5	PowerRadeRed Lipton	PowerRadeRed	FP	Hopt
	6	PowerRadeRed Mother	PowerRadeRed	FP	Coke
	7	Redbull RedbullSugar	Redbull	FP	Lipton
5b	1	E2Orange Hopt	E2Orange	P	
	2	LiftGreen Water	LiftGreen	FP	PowerRadePurple
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V Lift	V	P	
	5	Lipton PowerRadeRed	Lipton	FP	Hopt
	6	Mother CokeZero	Mother	FP	Coke
	7	Lipton Mother	Lipton	P	
5c	1	E2Orange Hopt	E2Orange	P	
	2	LiftGreen PowerRadePurple	LiftGreen	FP	PowerRadePurple
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V Lift	V	P	
	5	PowerRadeRed Lipton	PowerRadeRed	FP	Hopt
	6	VBlue Redbull	VBlue	FP	Coke
	7	LiftGreen Redbull	LiftGreen	FP	Lipton
5d	1	E2Orange Hopt	E2Orange	P	
	2	LiftGreen PowerRadePurple	LiftGreen	FP	PowerRadePurple
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V Lift	V	P	
	5	Hopt PowerRadeRed	Hopt	P	
	6	Mother LiftGreen	Mother	FP	Coke
	7	Redbull RedbullSugar	Redbull	FP	Lipton
5e	1	E2Orange Hopt	E2Orange	P	
	2	LiftGreen	LiftGreen	FP	PowerRadePurple

		Mother			
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V Lift	V	P	
	5	Lipton Hopt	Lipton	FP	Hopt
	6	LiftGreen Mother	LiftGreen	FP	Coke
	7	RedbullSugar LiftGreen	RedbullSugar	FP	Lipton

Using SURF in a controlled lighting environment

Image Name	Object Number	Histogram Top 2 results	Feature Extraction Result	Positive (P) /False Positive (FP)	The correct Result is
1a	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Coke Mother	Coke	FP	Lipton
	3	E2Lemon Mother	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	PowerRadeZero Lift	PowerRadeZero	FP	PowerRadeBlue
	7	RedbullSugar PowerRadeBlue	RedbullSugar	P	
1b	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Mother Water	Mother	FP	Lipton
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	PowerRadeZero Lift	PowerRadeZero	FP	PowerRadeBlue
	7	RedbullSugar PowerRadeBlue	RedbullSugar	P	
1c	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Water Coke	Water	FP	Lipton
	3	E2Lemon PowerRadeZero	E2Lemon	P	
	4	E2Orange E2Lemon	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	Lift PowerRadeZero	Lift	FP	PowerRadeZero
	7	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue

1d	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Water Coke	Water	FP	Lipton
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange E2Lemon	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	Lift PowerRadeZero	Lift	FP	PowerRadeZero
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
1e	1	PowerRadePurple Water	PowerRadePurple	P	
	2	Water Mother	Water	FP	Lipton
	3	E2Lemon E2Orange	E2Lemon	P	
	4	E2Orange Hopt	E2Orange	P	
	5	VBlue Redbull	VBlue	P	
	6	PowerRadeZero Lift	PowerRadeZero	P	
	7	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
2a	1	Coke Water	Coke	P	
	2	CokeZero V	CokeZero	FP	V
	3	CokeZero Mother	CokeZero	FP	Mother
	4	Lift PowerRadeZero	Lift	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	VBlue Redbull	VBlue	FP	RedBull
	7	GForce V	V	P	
2b	1	Coke Water	Coke	P	
	2	V GForce	V	P	
	3	Water Coke	Water	FP	Mother
	4	Lift PowerRadeZero	Lift	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	Redbull LiftGreen	Redbull	P	
	7	GForce V	GForce	P	
2c	1	Coke Water	Coke	P	
	2	V GForce	V	P	
	3	CokeZero Mother	CokeZero	FP	Mother
	4	Lift	PowerRadeZero	P	

		Mother			
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	Redbull LiftGreen	Redbull	P	
	7	GForce V	V	P	
2d	1	Coke Water	Coke	P	
	2	V LiftGreen	V	P	
	3	Water Coke	Water	FP	Mother
	4	Lift PowerRadeZero	Lift	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	Redbull LiftGreen	Redbull	FP	RedBull
	7	GForce V	GForce	P	
2e	1	Coke Water	Coke	P	
	2	V GForce	V	P	
	3	CokeZero Mother	CokeZero	FP	Mother
	4	Lift PowerRadeZero	Lift	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	VBlue Redbull	VBlue	FP	RedBull
	7	GForce V	GForce	P	
3a	1	Mother Water	Mother	FP	Hop
	2	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
	3	Water Mother	Water	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	RedbullSugar LiftGreen	RedbullSugar	P	
	7	Mother LiftGreen	Mother	FP	Water
3b	1	Coke Water	Coke	FP	Hopt
	2	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
	3	Water Coker	Water	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen Mother	LiftGreen	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	FP	RedBullSugar
	7	Mother PowerRadeRed	Mother	FP	Water

3c	1	Mother Water	Nother	FP	Hopt
	2	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
	3	Mother Water	Mother	P	
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	CokeZero LiftGreen	CokeZero	P	
	6	LiftGreen CokeZero	LiftGreen	FP	RedBullSugar
	7	Mother CokeZero	Mother	FP	Water
3d	1	Mother Coke	Mother	FP	Hopt
	2	RedbullSugar PowerRadeBlue	RedbullSugar	FP	PowerRadeBlue
	3	Coke Water	Coke	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	FP	RedBullSugar
	7	LiftGreen RedbullSugar	LiftGreen	FP	Water
3e	1	Water Mother	Water	FP	Hopt
	2	PowerRadeBlue RedbullSugar	PowerRadeBlue	P	
	3	Water Coke	Water	FP	Mother
	4	PowerRadeRed Lipton	PowerRadeRed	P	
	5	LiftGreen CokeZero	LiftGreen	FP	CokeZero
	6	LiftGreen RedbullSugar	LiftGreen	FP	RedBullSugar
	7	LiftGreen RedbullSugar	LiftGreen	FP	Water
4a	1	E2Lemon RedbullSugar	E2Lemon	P	
	2	LiftGreen Mother	LiftGreen	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Mother Water	Mother	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	LiftGreen RedbullSugar	LiftGreen	P	
	7	Mother Coke	Mother	FP	Lipton
4b	1	E2Lemon Lift	E2Lemon	P	
	2	Mother PowerRadeRed	Mother	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Coke	Coke	FP	Mother

		Water			
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hop
	6	LiftGreen Mother	LiftGreen	P	
	7	Lipton RedbullSugar	Lipton	P	
4c	1	E2Lemon Lift	E2Lemon	P	
	2	LiftGreen Mother	LiftGreen	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Mother Water	Mother	P	
	5	Mother PowerRadeRed	Mother	FP	Hopt
	6	LiftGreen CokeZero	LiftGreen	P	
	7	Mother Coke	Mother	FP	Lipton
4d	1	E2Lemon E2Orange	E2Lemon	P	
	2	Redbull RedbullSugar	Redbull	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Mother Water	Mother	P	
	5	Mother PowerRadeRed	Mother	FP	Hopt
	6	LiftGreen RedbullSugar	LiftGreen	P	
	7	Mother Coke	Mother	FP	Lipton
4e	1	E2Orange E2Lemon	E2Orange	FP	E2Lemon
	2	LiftGreen Redbull	LiftGreen	FP	Water
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	Coke Water	Coke	FP	Mother
	5	Mother PowerRadeRed	Mother	FP	Hopt
	6	LiftGreen CokeZero	LiftGreen	P	
	7	Lipton Water	Lipton	P	
5a	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Water	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V CokeZero	CokeZero	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke Water	Coke	P	
	7	PowerRadeRed Mother	PowerRadeRed	FP	Lipton

5b	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Coke	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V GForce	GForce	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke Water	Coke	P	
	7	Lipton PowerRadeRed	PowerRadeRed	P	
5c	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Water	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V GForce	V	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke PowerRadePurple	Coke	P	
	7	Water PowerRadeRed	Water	FP	Lipton
5d	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Water	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V GForce	V	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke Qater	Coke	P	
	7	PowerRadeRed Mother	PowerRadeRed	FP	Lipton
5e	1	E2Orange Hopt	E2Orange	P	
	2	PowerRadePurple Water	PowerRadePurple	P	
	3	PowerRadeZero Lift	PowerRadeZero	P	
	4	V GForce	GForce	P	
	5	PowerRadeRed Mother	PowerRadeRed	FP	Hopt
	6	Coke Water	Coke	P	
	7	Lipton PowerRadeRed	PowerRadeRed	P	